DOCUMENT RESUME

ED 326 466 SO 030 431

AUTHOR	Newmann, Fred M.
TITLE	The Relationship of Classroom Thoughtfulness to Students' Higher Order Thinking: Preliminary Results in High School Social Studies.
INSTITUTION	National Center on Effective Secondary Schools, Madison, WI.; Wisconsin Center for Education Research, Madison.
SPONS AGENCY	Office of Educational Research and Improvement (ED), Washington, DC.
PUB DATE	30 Jan 90
CONTRACT	G-00869007-89
NOTE	28p.
PUB TYPE	Reports - Research/Technical (143)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	Academic Ability; *Classroom Environment; Classroom Observation Techniques; *Classroom Research; Educational Research; High Schools; Skill Development; *Social Studies; *Student Evaluation; *Thinking Skills
IDENTIFIERS	*Classroom Thoughtfulness

ABSTRACT

An approach to the assessment of classroom thoughtfulness is presented. The approach recognizes the importance of in-depth knowledge, intellectual skills, and dispositions; it also emphasizes general qualities of discourse such as teachers posing higher order challenges and students giving reasons. The approach is contrasted with those that attempt to prescribe highly specific instructional procedures for teaching discrete thinking skills or specific bodies of content. The classroom observation scheme was used to assess levels of thoughtfulness in 39 diverse 9th grade social studies classes in 7 midwest high schools during the academic year 1988-1989. At the end of the year, students read two pages of background information on a constitutional issue and completed a written exercise asking them to state and to defend their position. Although teachers had not prepared students for such an exercise, the persuasiveness of student reasoning on the constitutional issues was strongly associated with the level of classroom thoughtfulness to which students were exposed, even after controlling for student scores on a pre-test of social studies knowledge, a pre-test of writing, student grade point average, race, sex, parents' education, and the racial and ability composition of the class. The design did not allow demonstration of a clear causal effect, but the evidence is consistent with the conclusion that general qualities of classroom discourse over a wide range of subjects affect student performance in higher order thinking. A 31-item bibliography is included. (Author/DB)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED326466

FINAL DELIVERABLE

The Relationship of Classroom Thoughtfulness to Students' Higher Order Thinking: Preliminary Results in High School Social Studies

Fred M. Newmann

January 30, 1990

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

This paper was prepared at the National Center on Effective Secondary Schools, supported by the U.S. Department of Education, Office of Educational Research and Improvement (Grant No. G-00869007-89) and by the Wisconsin Center for Education Research, School of Education, University of Wisconsin-Madison. The opinions expressed in this publication are those of the author and do not necessarily reflect the views of the supporting agencies. Thanks to Courtney Cazden, Adam Gamoran, Bruce King, James Ladwig, Joseph Onosko, Francis Schrag, James Shaver, and James Voss for helpful comments on earlier drafts.

SO 030 431

Abstract

An approach to the assessment of classroom thoughtfulness is summarized. It recognizes the importance of in-depth knowledge, intellectual skills, and dispositions, and it emphasizes general qualities of discourse such as students giving reasons and teachers posing higher order challenges. The approach is contrasted with those that attempt to prescribe highly specific teaching moves for teaching discrete thinking skills or specific bodies of content. The classroom observation scheme was used to assess levels of thoughtfulness in diverse social studies classes in seven high schools during an academic year. At the end of the year, students read two pages of background information on a Constitutional issue and completed a written exercise asking them to state and to defend their position. Although teachers had not prepared students for such an exercise, the persuasiveness of student reasoning on the Constitutional issue was strongly associated with the level of classroom thoughtfulness to which students were exposed, even after controlling for student scores on a pre-test of social studies knowledge, a pre-test of writing, student grade point average, race, sex, parents' education, and the racial and ability composition of the class. The design did not allow demonstration of a clear causal effect, but the evidence is consistent with the conclusion that general qualities of classroom discourse over a diverse range of subjects affect student performance in higher order thinking.

Researchers have repeatedly noticed the absence of thoughtful dialogue in classrooms (Cuban, 1984; Goodlad, 1984; Morrissett, 1982; Perrone & Associates, 1985; Powell, Farrar, & Cohen, 1985; Stake & Easley, 1978), and schools are flooded with diverse proposals to place more emphasis on the teaching of thinking (e.g. Costa, 1985; Marzano, Brandt, Hughes, Jones, Presseisen, Rankin, & Suhor, 1988; Pogrow, 1990; Sizer, 1984; Walsh and Paul, 1987). Some studies offer evidence that it is possible to improve students' thinking in certain ways with specific programs, but this research is often methodologically inadequate (Nickerson, 1988; Sternberg and Bana, 1986). And there has been virtually no research on the extent to which classroom thoughtfulness across teachers teaching a variety of classes without a common program for thinking affects student performance on a common task that calls for higher order thinking. We pursue this issue here by reporting on a new approach to the assessment of classroom thoughtfulness in high school social studies. The approach addresses critical issues in the research literature and offers a classroom observation instrument responsive to the practical needs of teachers. Initial results indicate that classroom thoughtfulness as assessed by the instrument is related to student performance. We begin by presenting the conception of higher order thinking and by showing how our approach addresses two central problems in the research literature on the teaching of thinking. This is followed by an account of an empirical study on the relationship of classroom thoughtfulness to student competence in reasoning on a civic issue.

I Critical Issues in the Conception and Teaching of Higher Order Thinking

Based on a review of philosophical, psychological and educational literature, we have defined higher order thinking as the interpretation, analysis, or manipulation of information to answer a question that cannot be resolved through the routine application of previously learned knowledge (Newmann, 1988). According to this definition, higher order thinking occurs whenever students respond to non-routine intellectual challenges. But the mere posing of higher order challenges offers no assurance that students will meet the challenges successfully. A useful pedagogical conception of thinking should identify the kinds of resources that students need to resolve higher order problems competently and what teachers can do to help students develop the resources. Consistent with other literature, we have explained elsewhere the need for three types of resources: in-depth knowledge, intellectual skills and dispositions of thoughtfulness (Newmann, 1990).

The main points of this perspective seem reasonably well accepted among researchers and informed practitioners (see, for example, Walsh and Paul, 1987). Controversy rages, however, over how to translate these general ideas into curriculum, pedagogy and assessment. Disagreement occurs on at least two levels: First, how much emphasis should be given to developing each of the three central resources - students' knowledge, skills and dispositions? Second, regardless of one's position on this issue, to what extent must knowledge, skills or dispositions - and pedagogies appropriate for each - be specified in detailed technical categories, as opposed to being conceived in more general, global terms? These issues can be summarized as the problem of priorities among central resources and

the problem of level of specificity. We discuss each of these problems and explain how our approach to assessing classroom thoughtfulness tries to resolve them in a way that is likely to advance practice.

A. Priorities Among Central Resources

Consider a teacher trying to help students answer the question, "Were the American colonists justified in using violence to secure their independence from England?" To enhance students' success in addressing this problem, how much attention should teachers give to developing students' knowledge, skills and dispositions? Building upon our previous review of literature (Newmann, 1990) we summarize here key arguments that can be made for each of these as the most critical resource.

The Knowledge Argument. Regardless of what side the student takes, a successful answer to this question demands in-depth knowledge of the circumstances of colonial life under British rule, including colonial grievances, British responses, principled arguments dealing with inalienable rights, taxation without representation, and ethical reasoning related to the destruction of property and the taking of human life. Beyond substantive knowledge about the historical period, students will need analytic knowledge; for example on elements of a well-reasoned argument, distinctions between empirical and normative issues, criteria for judging the reliability of evidence. Metacognitive knowledge may also be important, such as having a systematic approach for organizing one's thinking or an awareness of how one's thought processes and perceptions of others in a discussion might lead to error. The behavioral manifestations of some of these points might be labeled skills or dispositions, but they may all be considered knowledge in the sense that they all can be represented as cognitive beliefs. Skills and dispositions may facilitate the application of knowledge, but these points suggest that knowledge itself is the most critical foundation of understanding.[1]

The Skills Argument. Knowledge is undoubtedly important, but for the purposes of the teaching of thinking, skills are more critical, because they are the tools that permit knowledge to be used or applied to the solution of new problems. Some skills may be specific to the domain under study, and others more generic. To intelligently address the problem above, for example, one must be able to detect bias in the documents of colonial history and logical fallacies in inferences and arguments over the justification of the American revolution. One must be able to distinguish relevant from irrelevant information, to anticipate and to respond to arguments in opposition to one's own, to state one's views clearly and persuasively. Skills themselves may be construed or labeled in a variety of ways, but the main point is to recognize their role as cognitive processes that put knowledge to work in solving problems according to criteria for critical inquiry. In practice, knowledge is usually only transmitted from teacher to student without challenging the student to

[1]Various points in the argument for the centrality of knowledge have been made by Glaser (1984), McPeck (1981), and Nickerson (1988).

manipulate the knowledge to solve a higher order challenges. Unless the essential processes of using knowledge, i.e. skills, are stressed as central goals of education, higher order thinking is likely to be neglected and the knowledge transmitted to remain inert. Perhaps for this reason many educational reformers prefer not to advocate the teaching of thinking, but instead the teaching of thinking skills.[2]

The Dispositions Argument. Without dispositions of thoughtfulness, neither knowledge nor the tools for applying it are likely to be used intelligently. Those who argue for dispositions suggest several traits: a persistent desire that claims be supported by reasons (and that the reasons themselves be scrutinized); a tendency to be reflective - to take time to think problems through for oneself, rather than acting impulsively or automatically accepting the views of others; a curiosity to explore new questions, and the flexibility to entertain alternative and original solutions to problems. Thoughtfulness thereby involves attitudes, personality or character traits, general values and beliefs or epistemologies about the nature of knowledge (e.g., that rationality is desirable; that knowledge itself is socially constructed, subject to revision and often indeterminate; and that thinking can lead to the understanding and solution of problems). Knowledge and skills will be important for the mastery of particular challenges, but without dispositions of thoughtfulness, knowledge and skills are likely to be taught and applied mechanistically and nonsensically. Of the three main resources, dispositions have attracted the least attention in professional literature, but a good argument can be made that dispositions are central in generating both the will to think and in developing those artistic, ineffable qualities of judgment that steer knowledge and skills in productive directions.[3]

Our approach to the assessment of classroom thoughtfulness recognizes the legitimacy of each of the three resources, and we believe it is not possible to establish a defensible hierarchy among them. Thus, the observation scheme to be presented later is an attempt to capture the promotion of thoughtfulness through teachers' efforts to develop knowledge, skills and dispositions, without giving center stage to any one resource.[4] At the same time, we deliberately refrain from trying to assess the precise kinds of knowledge, skills and dispositions being promoted. The reasoning behind this choice relates to our conclusion on the next major issue.

B. Level of Specificity

---

[2]Various points in the argument for skills as the most central resource have been made by Beyer (1987), de Bono (1983), Herrnstein et al. (1986), Marzano et al. (1988).

[3]Various points in the argument for dispositions as a central resource have been made by Cornbleth (1985), Dewey (1933), and Schrag (1988).

[4]Those who emphasize interaction and interdependence among these resources include Bransford et al (1986), Ennis (1987), Greeno (1989), Perkins and Salomon (1989), and Walsh and Paul (1987).

The main issue here is the degree of precision and differentiation to strive for in identifying the kinds of knowledge, skills, and dispositions to guide instructional goals and pedagogy. At one end of the continuum is an orientation that strives toward ever increasing levels of specificity. At the other end is a perspective that strives for synthesis, integration, and holistic awareness.

Dominant approaches to curriculum, instruction, assessment and research itself seem to reflect a tendency toward increased specificity. Applied to our definition of higher order thinking, the conventionally accepted model for building curriculum and instruction can be summarized as four steps:

1. Identify the main problems or challenges that students should be competent to address (e.g., explanations of historical trends; developing positions on social issues; estimating and forecasting with sociological, economic or geographic data).

2. For each problem, identify the specific body of in-depth knowledge, the cluster of analytic skills, and the main dispositions needed for success in addressing the problem.

3. Experiment with alternative methods for teaching the specific knowledge, skills and dispositions relevant to each problem.

4. Codify the results to produce guidelines for curriculum and pedagogy most likely to assist students in resolving each of the major cognitive challenges identified in #1.

This approach seems systematic and reasonable, but it suffers from at least three potential inadequacies. First, by attempting to specify in advance the precise knowledge, skills and dispositions needed to solve particular problems and then teaching these directly, we risk over-programming students for success so that they may rarely have to confront novel challenges. The more practice one has in solving a particular type of problem, the more its mastery is likely to become routine and thus not involve a higher order challenge. Ironically, if carried to its extreme, this degree of programmed precision could actually reduce demands on the student for higher order thinking.

Second, such an approach escalates the specialization and balkanization of research into studies of countless specific topics and the many distinct types of knowledge, skills, and dispositions necessary for success on each. The large collection of experimental results for

how to teach a multiplicity of diverse problems would make it ever more difficult to synthesize findings useful for practitioners.[5]

Finally, by focusing exclusively on highly specific curriculum and pedagogy, this approach neglects several factors other than curriculum content that affect teachers' opportunities for success with students. At least three major influences on instruction are usually left untouched by studies of pedagogical precision: teachers' goals, philosophies and conceptions of knowledge; the kind of institutional leadership and organizational support given for higher order thinking; and characteristics of students that affect their degree of receptiveness to the promotion of thinking. Innovative specific teaching practices are surely needed, but improving teaching is far more complicated than discovering particular types of curriculum and pedagogy and then coaxing teachers to adopt them. Unless these additional factors are taken into account, there is little reason to believe that innovative specific pedagogy will be accepted, or, even if accepted, that it will significantly improve education.

This critique is not intended to suggest that we should always avoid programming students for success in specific tasks, that we should cease research on the teaching of specific topics, or that research on specific pedagogy will always be uninformative unless accompanied by research on broader issues of individual and institutional change. It is offered only to point out problems that have resulted from using the conventional model as the dominant approach to education research and development, without anticipating such consequences.

To minimize the problems raised by the dominant emphasis on highly specific lists of curriculum goals and pedagogical moves, we searched for another model. Rather than translating thinking into countless specific problems, skills and attitudes, we tried to identify more general qualities of classroom interaction that could be expected to help students face a variety of higher order challenges and that teachers would recognize as useful in advancing student thinking within the many domains of social studies. In other words, we began by asking what observable qualities of classroom discourse would be most likely to help students achieve depth of understanding, intellectual skills, and dispositions of thoughtfulness.

In theory, a classroom observation scheme might be more useful if derived from a validated model of how the mind learns and uses knowledge, skills and dispositions to solve problems. The point here, however, was not to articulate a general model of cognitive skills or to map the terrain of thinking processes that individuals follow as they work on

---

[5]The high degree of specificity that can occur in the naming of thinking skills is illustrated in Marzano et al. (1988) which notes twenty one different core thinking skills, including such items as defining problems, setting goals, observing, ordering, inferring, summarizing, establishing criteria.

problems. In spite of many advances in cognitive science, a validated model of cognitive process that accounts for success in meeting higher order challenges has yet to emerge, and so we refrained from endorsing any particular map of the terrain of thinking processes that individuals follow as they work on problems.

We believe that in-depth knowledge, intellectual skills, and dispositions of thoughtfulness can be developed through diverse specific activities, but also that certain general qualities of classroom interaction that nurture resources for thinking across a broad range of problems can be identified. Assessing general qualities of classroom discourse rather than highly differentiated behaviors helps to avoid fragmentation in teaching which itself can undermine student thinking. A more general, global approach may also hold more promise for transfer.

Our work with history and social studies teachers indicates that calls for specific types of thinking (e.g., critical, inductive, moral) are unlikely to generate widespread consensus for any particular type. Instead, social studies teachers are likely to perpetuate their previous emphases upon a plurality of types of thinking, but even these will be grounded primarily in the teaching of their subjects. Thus, a broad conception of thinking, adaptable to a variety of content and skill objectives, is more likely to generate serious interest among a diverse population of high school teachers.

A broad conception can strike at the heart of an underlying malady identified by many studies. At best, much classroom activity fails to challenge students to use their minds in any valuable ways; at worst, much classroom activity is nonsensical or mindless. The more serious problem therefore, is not the failure to teach some specific aspect of thinking, but the profound absence of thoughtfulness in classrooms. Even programs designed to teach thinking skills can fail to promote thoughtfulness. A general conception of thinking can address this basic issue.

II Indicators of Classroom Thoughtfulness

Here we present an observation scheme that recognizes the importance of all three resources (knowledge, skills and dispositions) and that minimizes the degree of differentiation in the assessment of teaching for thinking. The observation scheme was the main independent variable in the empirical study. In developing indicators responsive to the points made above, we used the following guidelines:

> The indicators should be able to be observed in the teaching of a variety of subject matter, skills, and dispositions within social studies.

> The indicators should refer to teacher behavior, to student behavior, and to activities involving both teacher and student.

> The indicators should allow for judgments on a continuum from less to more rather than merely discrete categorical values.

> The indicators should be conceptualized in ways that might later be used to help teachers reflect on their practice.

We rated lessons on more than fifteen possible dimensions of classroom thoughtfulness. After examining them from a theoretical point of view and with an awareness of some empirical qualities (distributions and correlations), we chose six as most fundamental.[6] Presumably, each of many dimensions represents a desirable characteristic that would contribute to thoughtful discourse. But there is an important distinction between a criterion for classroom discourse that indicates or helps to promote higher order thinking versus one that, in addition, seems so essential that one could not imagine judging a lesson "thoughtful" unless the criterion were met. Since we were not able to find analytic or empirical literature that conclusively justified a few key criteria, we put each of many dimensions to the following test: Based on the conception of higher order thinking outlined earlier, could a lesson conceivably score low on this dimension, yet still be considered a highly thoughtful lesson? If the answer was "yes," then the dimension was not considered critical or a minimal criterion. If the answer was "no," the dimension was judged as being minimally necessary, though perhaps not a sufficient, criterion for thoughtfulness.

The six main dimensions are described below. Each was used to make an overall rating of an observed lesson on a five point scale from 1 = "a very inaccurate" to 5 = "a very accurate" description of this lesson.

1. There was sustained examination of a few topics rather than superficial coverage of many.

Mastery of higher order challenges requires in-depth study and sustained concentration on a limited number of topics or questions. Lessons that cover a large number of topics give students only a vague familiarity or awareness and, thereby, reduce the possibilities for building the complex knowledge, skills and dispositions required to understand a topic.

2. The lesson displayed substantive coherence and continuity.

Intelligent progress on higher order challenges demands systematic inquiry that builds on relevant and accurate substantive knowledge in the field and that works toward the logical development and integration of ideas. In contrast, lessons that teach material as unrelated fragments of knowledge, without pulling them together, undermine such inquiry.

---

[6]The original complete list of indicators and illustrative reasoning on the selection of the final six are presented in Newmann (1990, in press).

3. Students were given an appropriate amount of time to think, that is, to prepare responses to questions.

Thinking takes time, but often recitation, discussion, and written assignments pressure students to make responses before they have had enough time to reflect. Promoting thoughtfulness, therefore, requires periods of silence where students can ponder the validity of alternative responses, develop more elaborate reasoning, and experience patient reflection.

4. The teacher asked challenging questions and/or structured challenging tasks (given the ability level and preparation of the students).

By our definition higher order thinking occurs only when students are faced with questions or tasks that demand analysis, interpretation, or manipulation of information; that is, non-routine mental work. In short, students must be faced with the challenge of how to use prior knowledge to gain new knowledge, rather than the task of merely retrieving prior knowledge.

5. The teacher was a model of thoughtfulness.

To help students succeed with higher order challenges, teachers themselves must model thoughtful dispositions as they teach. Of course, a thoughtful teacher would demonstrate many of the behaviors described above, but this scale is intended to capture a cluster of dispositions likely to be found in any thoughtful person. Key indicators include showing interest in students' ideas and in alternative approaches to problems; showing how he/she thought through a problem (rather than only the final answer); and acknowledging the difficulty of gaining a definitive understanding of problematic topics.

6. Students offered explanations and reasons for their conclusions.

The answers or solutions to higher order challenges are rarely self-evident. Their validity often rests on the quality of explanation or reasons given to support them. Therefore, beyond offering answers, students must also be able to produce explanations and reasons to support their conclusions.

The six dimensions were combined into a single scale (CHOT) which served as the indicator of classroom thoughtfulness for an observed lesson. As one part of a study on higher order thinking in high school social studies, four lessons from each of fifty one classes from grades 9 -12 were observed in 7 high schools, including courses as diverse as Introduction to Social Studies, US History, World History, Sociology, American Politics. Since a main point of this empirical study was to assess the relationship of classroom thoughtfulness to student performance, we turn next to a description of the dependent variable for student achievement.

III The Assessment of Students' Higher Order Thinking

To investigate the impact of classroom thoughtfulness on students' higher order thinking, we developed a social studies test which required organization, analysis, interpretation and manipulation of information. It was not feasible to devise a common test that also tapped the subject matter specific to each of the fifty one classes. Instead we created a task that represented an important civic issue which social studies should presumably equip students to think about and that provided sufficient information, not previously studied, for students to use.

A two-page document presented students with a court case involving the search of a student's (Karen) purse and locker by the high school principal who suspected Karen first of smoking in violation of a school rule and then of selling marijuana. Following the case description, background information was given on the main principles that courts have used in making decisions about the constitutionality of student searches. Students were asked to decide whether Karen's Constitutional rights were violated in the case and to write a persuasive essay which explained and defended their views by using information in the reading. Students' essays were scored from 1 to 5, based on criteria adapted from the assessment of persuasive writing of the National Assessment of Educational Progress (Applebee, Langer, Mullis, & Jenkins, 1990).

The dimensions of observed classroom thoughtfulness offer no information on how the teacher teaches persuasive writing, nor do they assess the nature of knowledge conveyed in class on the topic of the exercise (constitutional reasoning on searches in school). The dimensions were intended to identify general qualities of thoughtfulness rather than the quality of teaching to the specific demands of the student thinking task. The empirical question was whether general qualities of thoughtfulness would seem to promote competence in meeting specific cognitive challenges. To date, research has not systematically investigated this issue.

IV Methodology

A. Sampling of Schools, Classes and Lessons

Seven high schools in the midwest were selected to represent a diversity in social-economic composition. The schools were considered representative or typical in the sense that none were experiencing major reforms or dramatic changes during the period of data collection (1988-89). Demographic characteristics of the schools are given in Table 1. Within each school, several ninth-grade social studies classes were selected in order to maximize representation of diverse ability grouping patterns. The total number of ninth-grade classes

observed was thirty nine.[7] Each class was observed four times during the academic year (twice in the fall and twice in the spring). The lessons were selected randomly, based on scheduling convenience of the researchers and teachers, except for the provision that the lesson should include teacher participation (full-period films or test sessions were excluded). Teachers were aware of the general purpose of the study - to assess the promotion of higher order thinking, - but they had no knowledge of the specific dimensions of thoughtfulness that guided the observations.

B. Reliability of Classroom Thoughtfulness Indicators

Drawn from a three-person research team, different pairs of observers rated 24 lessons independently. Table 2 presents data on the extent of agreement between observers and correlations among the various pairs of ratings on each of the six dimensions. For most of the dimensions, there were high levels of inter-rater agreement.

In addition to inter-rater agreement, we examined the internal consistency of the CHOT scale. The Cronbach alpha of internal consistency among the six items (using the scores of all lessons observed) was .84. In addition, LISREL analysis indicated that the six dimensions when considered as one factor of thoughtfulness provided a better fit to the data than a model that specified thoughtfulness as 16 different dimensions on which the lessons were observed.

C. Test Administration, Scoring and Reliability

During a class period of approximately fifty minutes toward the end of the academic year, the test on student searches was administered to all classes. All students were able to complete the exercise during this period. The tests were scored from 1 to 5 by a team of 6 raters who developed specific content criteria to elaborate upon the general criteria for persuasive writing used in the National Assessment of Educational Progress. The general criteria are given in the Appendix. To determine inter-rater agreement, different pairs of two raters read 225 tests. The overall correlation was .76. Raters achieved exact agreement in 59% of the cases and agreed exactly or missed by only one point in 97% of the cases.[8]

---

[7]The ninth grade sample was part of a larger study of classes in grades 9-12, but since pre-test data were available only for ninth grade classes, this report is limited to those classes. Data presented below on reliability of classroom observations and scoring of the test of higher order thinking include classes above the ninth grade.

[8]The reliability rates in this study are somewhat lower than those achieved in the NAEP scoring of persuasive writing (Applebee et al. 1990), but this is to be expected, because our scoring required complicated judgments about students' use of subject matter content.

D. Other Variables in the Analysis

To assess the relationship of classroom thoughtfulness (measured by CHOT) to the test scores, we controlled for the influence of each of the following variables:

Student's grade point average (GPA) measured by the student's self report on an eight-point scale (1 = mostly below D, 2 = mostly D, 3 = about half C and half D, ... 8 = mostly A).

Student sex (MALE), measured by male = 1, female = 0.

Student race (AFAM), measured by African American = 1, other = 0.[9].

Parents' education (PAED), measured on a five point scale (1 = less than high school graduation, 2 = high school graduation only, ... 5 = graduate or professional degree) and averaged between two parents.

The average ability level of students in the class (CABIL), measured by the teacher's estimate of the percentage of students in the bottom third, middle third, and top third of the school's achievement distribution. Percentages were multiplied by 1, 2 or 3, and divided by 100 to yield a scale from 1 (low) to 3 (high).

Percentage of African American students in the class (CAFAM), according to teachers' reports.

Student pretest of social studies knowledge (NAPSCOR9). Administered in the fall of the academic year, this test consisted of multiple choice and short answer items drawn from previous NAEP tests of social studies, scored 0-79.

Student pretest of writing ability (ESSAY9). Administered in the fall of the academic year, this test asked students to write an essay (in 15 minutes) about a place or a possession that was important to them, to describe it "as fully as you can and explain why it is important to you." The test was scored from 0 - 9, based on the amount of information given and the level of abstraction.

V Results

Results are presented in Tables 3 - 5. The means and standard deviations in Table 3 indicate that performance on the higher order thinking task (SCORE) is low, with 68% of the students scoring 1 - 3 on the five-point scale. This confirms previous reports of low levels of student competence in writing about complex problems. Levels of classroom

[9]In this sample, most non-white students were African-American.

thoughtfulness (CHOT) also tend toward the lower end of the five-point scale, with most classes falling between 2.3 and 3.4, a finding consistent with other studies that describe the low levels of cognitive work done in high school classrooms. The ability level of the classes, students' grade point average, parents' education, and students' sex all seem to cluster around "average" values. The percentage of African-American students rounds to the national average of 13.7% for high school sophomores in 1980.[10] Scores on the pretests reflect mid-range values and adequate variance.

Examining the correlations in Table 4 we are most interested in variables associated with SCORE and CHOT. As expected, the test scores' strongest relationships are with the social studies pretest scores and the students' grade point averages. Note also, however, the size of the correlations of test score with classroom thoughtfulness, the ability level of students in the class, the writing pretest, and the percent of African-Americans in the class (negative relationship).

Other reports of this research will delve deeper into the possible determinants of classroom thoughtfulness (CHOT), by considering the degree of variance in thoughtfulness between teachers and schools and how this variance can be explained by individual and institutional characteristics. But the correlations indicate the possibility of a disturbing high negative relationship between classroom thoughtfulness and the percentage of African-American students in the class (-.42) which, in turn, may underlie the negative relationship ($r = -.29$) between that variable and scores on the civic reasoning test. Is this evidence of a racist tendency to deny opportunities for higher order thinking to classes with larger proportions of African-American students?

As might be expected we also find that classroom thoughtfulness scores increase somewhat with the ability level of the class, with parents' education, and with pretest scores on social studies knowledge. These findings can be explained by reasoning that teachers' expectations for student performance influence the degree to which they promote higher order thinking, that expectations themselves are determined largely by teacher assumptions about student ability, and that assumptions about ability are based in turn on teachers' perceptions of students' social background and previous success in school.

The regression presentation in Table 5 offers a more informative estimate of the relationship between classroom thoughtfulness and student test scores. Focusing on the

---

[10]From student self-report, the percentage of African American students is 11.9. CAFAM of 13.66% represents teachers' reports of the percentage of African American students enrolled in the class. The discrepancy may be due to a higher proportion of African Americans being absent or not identifying their race (as other studies have shown), or, if the student reports are more accurate, the sample containing a slightly lower percentage than the national average. Enrollments of other minority students were too low to consider in analysis.

standardized coefficients (Beta) and significance levels, we find four variables having the most robust relationships: social studies pretest knowledge, student's grade point average, classroom thoughtfulness, and writing competence on the pretest. That classroom thoughtfulness survives with a significant effect after controlling for all the other variables is a major discovery whose implications will be discussed below, but first we should note other findings in Table 5.

After controlling for all variables simultaneously, social background variables diminish in importance from what was implied in the bivariate correlations. That is, after actual achievement is taken into account (through individual gpa and pretest scores), the correlations (Beta) of race, ability level of the class, and parents' education with test scores disappear.[11]

The relatively large coefficient of the social studies pretest score (Beta = .22) deserves comment. Items on the pretest did not require students to synthesize or analyze information on constitutional issues, but the test did include many questions on government and the political process. Student success on the pretest may indicate not only the availability of background knowledge that might be applied to reasoning about constitutional issues, but also and perhaps more importantly, an interest in civic matters which leads to more competent performance when complex thinking in this area is called for.

VI Implications

Before discussing broader implications of the results, let us first look more specifically at the size of the relationship between classroom thoughtfulness (CHOT) and students' higher order thinking (SCORE). Using the raw regression coefficients (B) in Table 5, we see that increasing classroom thoughtfulness by one point on the five-point scale will, on average, with other variables held constant, lead to a fifth of a point increase in test scores (also on a five-point scale).[12] In comparison, to achieve this much gain on the thinking test, students would need an increase of one full letter grade in gpa (i.e. more than 2 points in an 8 - point letter grade scale) or an increase of 13 items correct on the social studies pretest. Considering that after taking all these variables into account, much of the variance

---

[11]Although females continue to perform slightly higher than males (significance level of .01), the magnitude of the difference (Beta = -.08) is too low to be educationally significant.

[12]Multinomial logit analysis indicated that increases in CHOT were significantly related to scores in the mid-range of the test score distribution. That is, they were associated with student differences between minimal and adequate and between adequate and elaborated, but did not seem related to differences between unsatisfactory and minimal or between elaborated and exemplary.

(about 72%) in individual higher order thinking performance is still unexplained, and that previous research when controlling for similar factors usually fails to find much influence of any instructional behavior, we view this finding as having potentially major significance. This should be qualified, however, by the realization that according to the standardized (Beta) coefficients, (a) the social studies pretest has a much greater association with the post test than does CHOT, and (b) none of the variables directly explains more than 5% of the total variance in test scores.[13]

We are in the process of replicating the study with an important modification; namely, use of a pretest similar to the test of higher order thinking. This is needed to determine the extent to which classroom thoughtfulness actually increases competence in the specific kind of civic reasoning we have chosen as the main dependent variable. As we await results of this work, we can, nevertheless, anticipate some implications of sustaining the present findings. The findings cast new light on the two main issues discussed earlier: priorities among the central resources that students need to use their minds well, and the degree of differentiation in the goals and processes of teaching that is likely to advance student performance in higher order thinking. thinking.

The relationship of classroom thoughtfulness to student writing on a civic issue offers new evidence on the question of whether competent higher order thinking is promoted by exclusive attention to domain-specific content, skills, dispositions of thoughtfulness or some combination of the three. The evidence here is not only consistent with the "combination" hypothesis. It is the most systematic quantitative test and the first demonstration of this relationship that we have seen. Since we have not done a comparative study between teachers that concentrate exclusively on in-depth content versus skills versus dispositions versus the various combinations, we cannot conclude that the full combination is the "best." It is possible, for example, that teachers who have the highest degrees of substantive and pedagogical knowledge in their content areas (as in the models of practice suggested by Shulman, 1987, or Wineburg and Wilson, 1988) would also score highest on classroom thoughtfulness. That is, possession of in-depth teaching competence in the subject area may necessarily express itself as promoting for students the skills and dispositions required to master higher order challenges. Given the consistent observation about the mindless kind of work that occurs in many social studies classes, it is, nevertheless, important to have progressed beyond rhetorical pleas for more rigorous teaching and to have shown that general qualities of thoughtfulness which include nurturing in-depth knowledge, skills and dispositions may indeed payoff in student performance on a complex task of civic reasoning.

On the question of differentiation, the results show that it is possible to identify some significant types of instructional interaction without having to count and code a large

---

[13]While much variance in individual test scores remains after the effect of CHOT has been estimated, we are most interested in the findings that on average, test scores do increase significantly with increases in CHOT.

number of highly specific teacher and student behaviors. Instead, it is possible to use a few general categories of interaction that characterize the lesson as a whole and that take into account the nature of the lesson's content, the reactions of students and the teacher's overall teaching style. The relative simplicity of our observational categories has two main advantages. First, it resists the reductionistic tendency to dissect the process of teaching into so many discrete units that teaching itself turns into a fragmented, mindless enterprise. Second, it offers a set of criteria which teachers could use to reflect upon their practice without a great deal of technical training or time consuming analysis.

Having identified instructional qualities related to student competence in higher order thinking in high school social studies, much work remains. The power of the observational scheme must be further tested by refinement of the pretest. If these findings are sustained, it will be even more useful to concentrate on the question of how to produce higher levels of classroom thoughtfulness. Research on this issue has been underway since 1986. Future reports will seek to explain the extent to which levels of classroom thoughtfulness can be enhanced within a department through specific policies, programs and approaches to leadership by high school principals, department heads and teachers.

References

Applebee, A. M., Langer, J. A., Mullis, I. V. S., & Jenkins, L. B. (1990). The writing report card, 1984-1988: Findings from the nation's report card. Princeton, NJ: National Assessment of Educational Progress.

Beyer, B. (1987). Practical strategies for the teaching of thinking. Boston: Allyn & Bacon.

Bransford, J., Sherwood, R., Vye, N., & Rieser, J. (1986). Teaching thinking and problem solving. American Psychologist, 41(10), 1078-1089.

Costa, A. (Ed.). (1985). Developing minds: A resource book for teaching thinking. Alexandria, VA: Association for Supervision and Curriculum Development.

Cornbleth, C. (1985). Critical thinking and cognitive processes. In W. B. Stanley (Ed.), Review of research in social studies education: 1976-1983. Boulder, CO: ERIC Clearinghouse for Social Studies/Social Science Education.

Cuban, L. (1984). Policy and research dilemmas in the teaching of reasoning: Unplanned designs. Review of Educational Research, 54(4), 655-681.

de Bono, E. (1983). The direct teaching of thinking as a skill. Phi Delta Kappan, 64(10), 703-708.

Dewey, J. (1933). How we think. Boston: D.C. Heath.

Ennis, R. H. (1987). A taxonomy of critical thinking dispositions and abilities. In J.B. Baron and R.J. Sternberg (Eds.), Teaching for thinking. New York: Freeman.

Glaser, R. (1984). Education and thinking: The role of knowledge. American Psychologist, 39, 93-1104.

Goodlad, J. I. (1984). A place called school: Prospects for the future. New York: McGraw Hill.

Greeno, J. G. (1989). A perspective on thinking. American Psychologist, 44(2), 134-141.

Herrnstein, R. J., Nickerson, R. S., De Sanchez, M., & Swets, J. A. (1986). Teaching thinking skills. American Psychologist, 41(11), 1279-1289.

Marzano, R. J., Brandt, R. S., Hughes, C. S., Jones, B. F., Presseisen, B. Z., Rankin, S. C. & Suhor, C. (1988). Dimensions of thinking: A framework for curriculum and instruction. Alexandria, VA: Association for Supervision and Curriculum Development.

McPeck, J. E. (1981). Critical thinking and education. New York: St. Martins.

Morrissett, I. (Ed.). (1982). Social studies in the 1980s: A report of project SPAN. Alexandria, VA: Association for Supervision and Curriculum Development.

Newmann, F. M. (1990). Higher Order Thinking in Teaching Social Studies: A Rationale for the Assessment of Classroom Thoughtfulness. Journal of Curriculum Studies, 22(1), 41-56.

Newmann, F. M. (in press). Qualities of thoughtful social studies classes: An empirical profile. Journal of Curriculum Studies.

Newmann, F. M. (1988, May). Higher order thinking in the high school curriculum. NASSP Bulletin, 72(508), 58-64.

Nickerson, R. S. (1988). On improving thinking through instruction. In E.Z. Rothkopf (Ed.) Review of Research in Education, 15, (pp. 3-57). Washington, DC: American Educational Research Association.

Perkins, D. N., & Salomon, G. (1989). Are cognitive skills context-bound? Educational Researcher, 18(1), 16-25.

Perrone, V. & Associates. (1985). Portraits of high schools: A supplement to high school: A report on secondary education in America. Princeton, NJ: Carnegie Foundation for the Advancement of Teaching.

Pogrow, S. (1990). Challenging at-risk students: Findings for the HOTS program. Phi Delta Kappan, 71(5), 389-397.

Powell, A. G., Farrar, E., & Cohen, D. K. (1985). The shopping mall high school: Winners and losers in the educational marketplace. Boston: Houghton Mifflin.

Schrag, F. (1988). Thinking in school and society. London: Routledge & Kegan Paul.

Shulman, L. S. (1987, February). Knowledge and teaching: Foundations of the new reform. Harvard Educational Review, 57(1), 1-22.

Sizer, T. R. (1984). Horace's compromise: The dilemma of the American high school. Boston: Houghton Mifflin.

Stake, R. E., & Easley, J. A. (1978). Case studies in science education. Washington, DC: National Science Foundation.

Sternberg, R.J. and Bhana, K. (1986). Synthesis of research on the effectiveness of intellectual skills programs: Snake-oil remedies or miracle cures? Educational Leadership, 44(2), 60-67.

Walsh, D., & Paul, R. W. (1987). The goal of critical thinking: From educational ideal to educational reality. Washington, DC: American Federation of Teachers.

Wineburg, S. S., & Wilson, S. M. (1988). Models of wisdom in the teaching of history. Phi Delta Kappan, 70(1), 50-58.

Appendix

Criteria for Scoring of Persuasive Writing

OVERVIEW

Students' essays will receive one of five grades: (1) unsatisfactory, (2) minimal, (3) adequate, (4) elaborated, or (5) exemplary. The overarching consideration is the degree to which a student's response is capable of persuading a reader. Three elements will focus this assessment: whether or not the student has a) taken an informed stand, b) provided persuasive reasons, and c) elaborated upon those reasons. Specific points will not be subtracted for unpersuasive or irrelevant reasons but these could diminish persuasiveness. In addition, presentation of undermining reasons or faulty assumptions (with respect to the text of the test only) can also diminish persuasiveness. Finally, responses should be written in full sentences; that is, incomplete sentences or fragmented lists are considered less persuasive. Descriptions of the five types of responses are provided below to serve as a scoring guide when grading essays.

1. Unsatisfactory: The student has failed to take a stand on the issue under examination, or has taken a stand but has failed to provide a single persuasive reason. Lacking a persuasive reason, unsatisfactory responses will necessarily lack elaboration. Overall, the response has no chance of persuading the reader.

2. Minimal: The student has taken a stand on the issue under examination and has provided at least one persuasive reason, or at least two supportive reasons. Faulty assumptions, undermining, or irrelevant reasons could result in an unsatisfactory score if they reduce the persuasiveness of the argument. Overall, however, the response is unlikely to persuade the reader.

3. Adequate: The student has taken a stand and has provided two or more persuasive reasons. Elaboration of reasons is not necessary here. The presentation of only one persuasive reason can result in a score of "adequate" if useful elaboration is included. Undermining reasons, faulty assumptions, or irrelevant reasons can possibly reduce the score to "minimal." Overall, the response has a chance of persuading the reader.

4. Elaborated: The student has taken a stand, has provided two or more persuasive reasons, and has provided elaboration on at least one of those reasons. Presentation of many persuasive reasons (at least 3) without elaboration can also produce this score. Undermining reasons, faulty assumptions, or irrelevant reasons can possibly reduce the score. Overall, the response is likely to persuade the reader.

5. Exemplary: The student's response meets criteria for (4) above, and demonstrates (a) at least two elaborated persuasive reasons, and (b) an argument so clear and coherent (i.e., no significant undermining reasons, faulty assumptions or irrelevant reasons) and grammatically correct as to merit public display as an outstanding accomplishment for a high school student. Overall, the response is more likely to persuade the reader.

Table 1

Demographic Profiles for Representative Schools*

	SCHOOL						
	21	22	25	26	27	28	29
1. 1988 Enrollment	1385	768	1807	430	1637	1447	1740
2. Ethnic Racial Composition							
a) % White	76	99	95	95	37	50	80
b) % African American	18	0	2	1	27	32	9
c) % Asian	4	0	1	2	3	4	8
d) % Hispanic	2	0	2	1	30	15	3
e) % Other Minority	0	0	0	1	4	0	0
3. Low Income	0	8	20	5	30	16	11
4. Number of Teachers	103	59	102	35	110	108	111
5. % 1988 Graduates Going To:							
a) % 4 Year College	82	48	55	76	31	25	82
b) % Technical School	3	24	1	8	7	5	4
c) % 2 Year Community College	3	0	15	8	34	32	0
d) % Military	2	3	5	2	8	8	1
e) % Job + Other	10	23	24	6	20	30	13
6. 1988 Percent Drop-out Rate, Based on 4 Years	5	4	0	0	12	7	2
7. Per Pupil Expenditure 1988	7163	3785	2616	3181	4100	3374	4700

* Because all percentages are rounded to nearest whole number, they may not add to 100%.

Table 2

Inter-Rater Agreement for Dimensions of Th[redacted]ness (Scored 1-5) on 24 Lessons

Dimension	% Exact Agre[redacted]	Differ by 1 Point or Less	r**
1. Few Topics	70.8	100.0	.88
2. Coherence	75.0	100.0	.86
3. Enough Time	66.7	91.7	.50
4. Cognitive Challenge	58.3	100.0	.87
5. Teacher Models Thoughtfulness	69.6	91.3	.78
6. Student Reasons & Explanations	62.5	91.7	.84

*% agreement is based on 24 ratings per dimension by two raters.
**Pearson correlation based on 3 different pairs of raters for each dimension across the 24 lessons.

Table 3

Means and Standard Deviations of Test Scores, Classroom Thoughtfulness and Background Variables, 9th Grade Students

	MEAN	STD DEV	CASES
SCORE	2.04	.91	734
CHOT	2.83	.56	734
CAFAM	13.66%	13.29	734
CABIL	1.99	.50	734
AFAM	12%	---	723
MALE	47%	---	724
PAED	3.32	1.04	649
GPA	5.57	1.53	727
NAPSCOR9	57.46	12.84	734
ESSAY9	5.51	1.35	719

SCORE = Post-Test Persuasive Constitutional Essay, 1-5
CHOT = Classroom Thoughtfulness, 1-5
CAFAM = % of African American Students in Class
CABIL = Average Ability Level of Class, 1-3
AFAM = Student's Race, African American = 1, Other = 0
MALE = Student's Sex, Male = 1, Female = 0
PAED = Parents' Education Attainment, 1-5
GPA = Student's Grade Point Average, 1-8
NAPSCOR9 = Pre-Test Social Studies Knowledge, 0-79
ESSAY9 = Pre-Test Writing Abililty, 0-9

Table 4
Correlations Among Variables

	SCORE	CHOT	CAFAM	CABIL	AFAM	MALE	PAED	GPA	NAPSCORE9	ESSAY9
SCORE	1.00									
CHOT	.31	1.00								
CAFAM	-.29	-.42	1.00							
CABIL	.30	.25	-.44	1.00						
AFAM	-.20	-.18	.39	-.16	1.00					
MALE	-.13	-.12	.09	-.09	.06	1.00				
PAED	.22	.23	-.17	.13	.01	-.03	1.00			
GPA	.36	.18	-.19	.29	-.10	-.08	.27	1.00		
NAPSCOR9	.43	.31	-.35	.41	-.30	.09	.25	.44	1.00	
ESSAY9	.29	.14	-.11	.24	-.08	-.18	.19	.26	.29	1.00

N of cases encountered = 734

Minimum Pairwise N of Cases = 643

Table 5

Regression of Test Score on Classroom Thougthfulness and Background Variables

Multiple R	.54
R Square	.29
Adjusted R Square	.28
Standard Error	.77

Analysis of Variance

	DF	Sum of Squares	Mean Square
Regression	9	174.89	19.43
Residual	724	426.05	.59

F = 33.02 Signif F = .00

Variable	B	SE B	Beta	T	Sig T
GPA	.09	.02	.16	4.39	.00
MALE	-.15	.06	-.08	-2.51	.01
AFAM	-.14	.10	-.05	-1.46	.14
CHOT	.21	.06	.13	3.59	.00
PAED	.04	.03	.05	1.34	.18
CABIL	.11	.07	.06	1.62	.11
CAFAM	-.00	.00	-.05	-1.30	.19
NAPSCOR9	.02	.00	.22	5.48	.00
ESSAY9	.08	.02	.12	3.55	.00
(Constant)	-.63	.25		-2.53	.01